More and more of our personal health data are being stored electronically. And as that information piles up, researchers are mining it for answers that can help guide research and improve health care.

This is a report for iHealthBeat, a daily news service of the California HealthCare Foundation. I'm Erika Kelly.

You know that traditional image of a doctor sitting with a patient, pen in hand, scribbling on a chart? These days, your doctor or nurse is more likely to be sitting at a keyboard.

(Sound of Typing)

They're typing up notes and checking boxes, entering data into your electronic health record.

(Selby): "Data such as your weight, or your height, your blood pressure."

And that's not all, says Joe Selby, executive director of PCORI, the Patient-Centered Outcomes Research Institute.

(Selby): "Lab results are entered. The results of tests like X-rays. What kinds of medications were prescribed. What diagnoses were considered, and what kinds of diagnoses were made."

Health care providers across the country are adopting EHRs, motivated in part by incentives from the federal government. Selby says they're a rich source of clinical information that can help answer the questions that are most important to patients.

(Selby): "For example, about which one of these two medications that I'm considering will control my symptoms better and cause the fewest side effects?"

To help answer those questions more quickly, PCORI is spending millions of dollars to develop PCORnet, a national health data network. It includes clinical data research networks based at hospitals and health care centers. It also includes patient-powered networks, which engage people with particular conditions. Selby says the goal is for research and health care to go hand in hand. To find out how treatments work in the real world.
(Selby): "We'll get more accustomed as a country to asking and answering a much larger number of questions, using a larger amount of data. We'll be able to focus in on the differences between patients and really get to that point where we can ask, 'What works best for whom?'"

But mining EHRs for answers is not a simple task. One big hurdle? Health care providers use many different EHR systems, and clinicians have different ways of entering data. Selby says it's like they're speaking different languages. That's why PCORnet members have to translate their data into a common data model. Think of it like a decoder ring, says Elizabeth McGlynn. She's the director of the Kaiser Permanente Center for Effectiveness and Safety Research.

(McGlynn): "It's a way of saying, there may be 27 ways that somebody could define somebody as having diabetes, and if any of those are true, and algorithms can certainly be more complicated than that, then we're going to take that and code it as diabetes in the common data model."

McGlynn says it's a big job, which takes time and money. But, she says, it helps make sure that all that data are rested and ready -- whether they're collected at Kaiser in California, or some place across the country.

(McGlynn): "So the work of the common data model is to create these standard definitions so our ability to use those data for research across these very different places is made much easier."

PCORnet -- which is still in its early stages -- is not the only effort to draw research from EHRs. Nigam Shah is a professor of bioinformatics at Stanford University. He and his colleagues are applying data mining to several different research areas, including pharmacovigilance.

(Shah): "In pharmacovigilance, you can find associations between drugs and adverse events."

Shah co-authored a recent study that found an association between proton pump inhibitors, like Nexium and Prilosec, and an elevated risk of heart attack. Using a novel data mining technique, they were able to comb through millions of EHRs, including what's known as unstructured data.

(Shah): "So the novelty here is to be able to use the clinical documents, the admission notes, the discharge notes and those kinds of things."

Shah says many research studies that draw on EHRs only focus on the structured data. Information that's gathered from the checkboxes and drop-down menus that are common parts of EHRs. But Shah says they're missing out.
(Shah): "If you ask a group of clinicians, 'What portion of the useful information is in the unstructured data?' You get answers ranging from 50% to 95%.

Still, Shah emphasized that their data mining research does not prove causation between proton pump inhibitors and heart attacks. Instead, he says, it can help prioritize drugs for further safety research. That's an idea echoed by Susan Weber, director of Informatics Services at the Stanford Center for Clinical Informatics.

(Weber): "I think data mining has potential for pointing researchers into directions that they might not have otherwise considered exploring. Ultimately, you're going to have to validate the results of the data mining research studies with structured, prospective, double blinded, conventional clinical studies."

Weber and her team help researchers like Shah get access to Stanford's vast database of EHRs. But first they take steps to protect patient privacy. For some studies, that means making sure that data are de-identified.

Stanford, Kaiser and other health care centers also rely on secure servers, firewalls and other cybersecurity measures to protect patient data. Still, they acknowledge that nothing is foolproof. Health information, like all data, is vulnerable to hacks and leaks. But they argue that the potential of this kind of research is worth the risk.

(Weber): "The community is aware of that, so we're all very cautious. In the end, it's a matter of pragmatics and practicality. Is someone really going to be maliciously trying to target individuals in this way, or is it OK to continue doing research for the betterment of mankind?"

This has been a report for iHealthBeat, a daily news service from the California HealthCare Foundation. If you have feedback or other issues you'd like to have addressed, please email us at IHB@chcf.org. I'm Erika Kelly. Thanks for listening.